

H. Varga Gyula

A magyar nyelv prefixum típusú elemei

(Formális és automatikus leírás)

Kandidátusi értekezés

Magyar Tudományos Akadémia

Budapest, 1993

1. Az értekezés tárgya, célja, tudományos előzményei

Amikor az ember kommunikál, közlendőinek átadhatóvá tételéhez – meglepő könnyedséggel és sebességgel – rendkívül bonyolult (agyi-logikai) műveletek sorozatának a végén egy, számára adott invertáriumból (lexikonból) különféle elemeket válogat össze valamiféle szabályossággal. A logikai műveleteket is, az alapelemeket is, az összeszerkesztés szabályait is évszázadok óta fürkészi a gondolkodó ember, mégis úgy tűnik, minden generációnak marad további kutatnivaló.

Több mint száz éve, Saussure óta (úgy) tudjuk, hogy a nyelvi készlet alapegységei az alakkal és jelentéssel bíró jelek, a morféma. Azt is tudjuk, hogy ezek – nyelvi szerepüket tekintve – két nagy típusba sorolhatók: főelemek és viszonyító- vagy mellékelemek. Ez utóbbiak (névelők, névutók, elöljárók) egy része viszonylag szabadon „kószál” a főelemek (szótövek) környékén, más része viszont (a toldalékok vagy affixumok) hozzáragadt a szótövhöz. A szótő afféle supermorféma: a világ minden nyelvében megvan, funkciójuk megegyezik vagy hasonló, formai viselkedésük nyelvtípusonként eltér. A viszonyítóelemek rendkívül differenciált állományából egyes nyelvek az egyik, mások a másik típust részesítik előnyben. Az affixumoknak többféle rendszerezése létezik, bennünket kiindulásként a szóalakban elfoglalt helyzete érdekel. Eszerint – a szótövhöz viszonyítva – beszélünk az azt megelőző (prefixum), beleékelődő (infixum) és követő (suffixum) toldalékokról.

A magyar mint jellegzetesen agglutináló nyelv a kutatók és az iskolai nyelvtanok szerint a toldalékok közül csak a suffixumokat használja. Ezzel a felfogással kíván vitatkozni a jelen dolgozat. Nyelvünkben ugyanis kimu-

tathatók olyan elemcsoportok, amelyek prefixumnak minősülnek. Közülük a legjelentősebb az igekötők szövevényes, folyamatosan gazdagodó, funkcionálisan és szemantikailag egyre differenciálódó állománya.

A dolgozat célja, hogy földerítse a magyar nyelv prefixum típusú elemeit, áttekintse azok főbb jellemzőit, rendszerét; új szempontok figyelembevételével minősítse az igekötőket, megadja formális leírásukat, a számítógépes nyelvészet számára pedig elvégezze az igekötős szerkezetek morfológiai és szintagmatikus alakjainak az automatikus analízisét és szintézisét.

A szakirodalmi előzmények

A mai magyar szakirodalom az alaktani elemek között a prefixumok csoportját nem tárgyalja, a grammatikai leírások ilyen morfológiátípust nem különítenek el. Ismeretes, hogy régi nyelvtanaink terminológiájára és szemléletére eleinte a latin, majd a német minta hatott erőteljesen, ilyenformán – az említett nyelvek kategóriáit keresve – a magyar nyelv jelenségeit is nemritkán amazok értékrendje szerint ítélték meg. Ezért találunk bennük ilyen – elsősorban az igekötő megnevezésére szolgáló – terminusokat: *prefixum*, *praepositio*, *elől-járó*, *előrag*, *előképző*.

Más nyelvek leírásából ismerjük ezt a morfológiátípust. Az orosz nyelv például a prefixumok óriási gazdagságát mutatja. Nemcsak igéket és főneveket, hanem mellékneveket és határozószókat is lehet alkotni prefixumokkal. A prefixáció mellett ismeretes a prefixo-suffixáció és a prefixo-posztfixáció, sőt a prefixo-suffixo-posztfixáció is (Svedova 1970: 253–293). Köztük a – nyelvi szerepük szerint talán legbonyolultabb – igei *prisztavkának* három fő funkciója különíthető el: szemantikai-szóképző, szintaktikai és aspektuális.

A németben közismert a nagyszámú *Vorsilbe*, melyek többsége ráadásul – bár nem olyan rend alapján, mint a magyar igekötő – szabály szerint el is válik igéjétől. A Duden prefixumokat (*be-*, *ent-*; *ge-*, *ver-* usw.) és félprefixumokat (Halbpräfix, z. B.: *aus-*, *über-*, *wieder-*) különít el. Fleischer az előbbieket homonim szabad morféma nélküli prefixumoknak, az utóbbiakat pedig homonim szabad morfémasáknak nevezi (Fleischer 1971: 295). A félprefixumok alakilag egybeesnek a határozószókkal, prepozíciókkal, jelentésbelileg azonban a prefixumokhoz állnak közel. (I. h.)

Az angol nyelvtanokban jóval kisebb szerepe van a morfológiának, mert a grammatikai jelentéseket az angol nyelv elsősorban analitikusan fejezi ki. Az *affixáción* belül azonban ott is fontos helyet kap a prefixáció. A Quirk-féle nyelvtan a prefixumok 10 csoportját sorolja föl (negative, reversative, locative etc. prefix). Az igei prefixumok közismertek (*dislocate*, *prepare*, *resurch*), az utolsó csoportbeli ún. konverzív prefixumok (conversion prefixes) főnévből képeznek igét, pl.: *bewich*, *befriend*, *embody*, *endanger* etc. (Quirk 1985: 1546).

Ha távolabb is akarunk tájékozódni, érdekességként megemlíthetjük, hogy például a csukcs-kamcsatkai nyelvekben – ezek mindegyike agglutináló – nemcsak a fő végéhez járuló szuffixumok használatosak, hanem a prefixumok is, azaz a fő elé járulók. (Lásd Antal 1970: 194.) Az afrikai nyelvek túlnyomóan szintén ragozók, de például a bantu nyelvekben az előragok (prefixumok) száma meghaladja az utóragokét. Sőt a főnevek típusokba sorolása is az előragok alapján történik. (Vö. Fodor 1970: 256.) A távolkeleti mon-khmer nyelvekben más a helyzet: ott szuffixumokat alig ismernek, a szóképzés egész rendszere a prefixumok alkalmazásán nyugszik (Antal 1970: 238).

Visszatérve a magyar nyelvhez, a grammatikai leírások a kötött morféma-kapcsolódási rendjéről megállapítják, hogy helyük a szóalakban a szótő mögött van. A képzők között nem tartanak számon prefixumokat. A toldalékok sorrendi szabályainak rendszerszerű vizsgálatával kapcsolatban elsőként Antal Lászlót (1961 és 1964) kell megemlítenünk. Disztribúciós elemzésében nem esik szó prefixumokról, de egyértelműen kiderül, hogy ilyen morfématípust nem tart számon a magyar nyelvben. Mind a képzők, mind a jelek ismérvei között szerepel, hogy „nem állhatnak soha szó eleji helyzetben” (Antal 1961: 68, ill. 75). Ebből következtethetünk arra, hogy az igekötős ige összetételnek tartja. S nem meglepő az sem, hogy a felsőfok jelével nem tud mit kezdeni.³

Lényegében ezen az állásponton van Berrár Jolán is a szóképzés vizsgálatának új szempontjait és módszereit tárgyaló tanulmányában (1973: 119).

Elszórva, a nyelvi elemek leírásában azonban mégis találkozunk a *prefixum* megjelöléssel: a felsőfok *leg-* elemének és az igekötőknek a minősítésében. A magyar nyelv története c. egyetemi tankönyv szerint toldalékaink mind szuffixumok – kivéve a *leg-* „szócskát” mint prefixumot, továbbá az önálló szavak és a toldalékok között átmenetet képező igekötőket (i. m. 213). Az akadémiai nyelvtannak a szóelemek általános kérdéseit taglaló fejezetében nem esik szó prefixumról, csupán a felsőfokjel kivételes pozíciójáról (Tompai 1961: 301). A fokozásról szóló részben a felsőfok jelét már egyértelműen prefixumnak minősíti. De az igekötő definíciója is így kezdődik: „Az igekötő (prefixum verbale)...” (i. m. 263).

Deme László a vizsgálandó morfématípust a toldalékok rendszerében jellemzi, s az affixumokat a szóalakban elfoglalt helyzetük alapján különbözteti meg (prefixum, infixum és szuffixum). Megjegyzi, hogy a magyarban a szuffixum a legismertebb; de prefixumszerű eszközünk is van: a *leg-*; s

³ „A *leg-* morfémanak nincs helye a jelek, illetőleg általában a toldalékok között, mert annyira elűt tőlük megjelenési helyét illetően” (Antal 1961: 75).

e felé van fejlődőben az igekötő is, ám hozzáteszi, hogy ebben erősen hátráltatja az, hogy elválhat az igétől (Deme 1976, 44).

Károly Sándor (1980: 121–156) – a morfológiai tagoltság változásait vizsgálva – két morfémátípust különít el: fő- és segédelemeket. Az előbbi fogalomszók és névmások alkotják (ez tkp. megfelel Deme szabad morfémáinak, az pedig Bloomfield rendszerezését sejteti), az utóbbit az affixumok és segédszók. Az affixumok között található az általunk vizsgált elemtípus is: prefixum: *leg(szebb)* (i. m. 125). Az igekötővel azonban a segédszavak egyik alcsoportjában találkozunk: „pre-postpositumok: *meg(fog)* – *(fog) meg*” (uo.).

A prefixummá minősítésben Gaál Edit megy legtovább. Egyik tanulmányában a kötött morfémák közé a prefixumot is besorolja: „Az affixumok előfordulhatnak szó eleji helyzetben is, ilyen a magyarban a *leg-* és a *legesleg-*. [...] Így lesz például a *leg-* jel prefixum” (Gaál 1991: 12–13).

2. A vizsgálati eredmények összefoglalása

A dolgozat három tartalmi egységből áll. Az első rész a prefixum fogalmát, jellemzőit és a morfémarendszerben való elhelyezését taglalja. Ezt követően kerül sor a magyar nyelvben föllelhető prefixum típusú elemek rendszerszerű áttekintésére. A második fejezet teljes egészében az igekötőkkel foglalkozik. Ez a rész az utolsó fejezet automatikus leírását készíti elő. Ehhez szükséges az igekötők különféle tulajdonságainak bemutatása. A minősítés és az állománykijelölés statisztikai vizsgálat alapján történik. A morfológiai jellemzés az igekötős lexémák alkotási eljárásait írja le, a szórenddel foglalkozó rész pedig az igekötős szerkezeteknek az eddigiektől eltérő szórendi tipizálását ajánlja. Az utolsó fejezet a számítógépes nyelvészet módszereivel végzi el az igekötős szóalakok analízisét és a szerkezetek szintézisét, végül bemutatja az alkalmazási lehetőségeket.

A különféle vizsgálatokhoz szolgáló korpuszként öt műfajból vett szövegrészletnek összesen mintegy 150 000 szövegszót tartalmazó állománya szolgált. A statisztikai feldolgozás egyik része az ELTE, másik a KLTE számítóközpontjában történt. A morfológiai és a szintaktikai programok – a dolgozatban található leírás és folyamatábrák alapján, Pascal nyelvben – az EKTf informatikai tanszékén készültek.

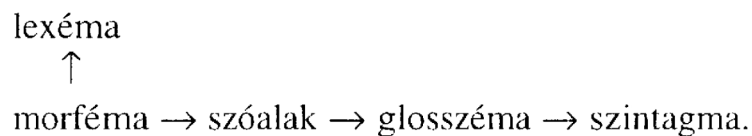
2.1. Prefixumok a magyar nyelvben

A magyar nyelvtanokban régóta elterjedt módszer, hogy a szófaji rendszer bemutatása mellett, attól viszonylag függetlenül kerül sor a szóelemek vizsgálatára. Az előbbi a szófajtan, az utóbbit a (szó)alaktan (morfológia) végzi. Az európai grammatikák másképp értelmezik a morfológiát. E címszó alatt tárgyalják a szó szerkezeti elemeit, azok nyelvtani szerepét, valamint a

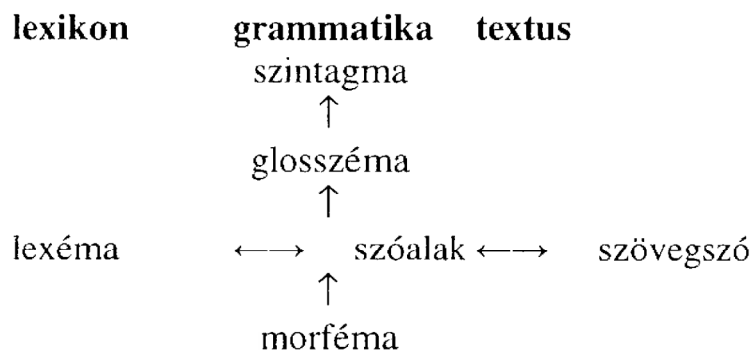
szófaji rendszert, beleolvastva a formaképzést (vagyis a paradigmaticát, tkp. a ragozást). A szóalkotást vagy a morfológia külön fejezeteként, vagy a morfológián kívül tárgyalják.

Mint láttuk, a mai magyar szakirodalom az alaktani elemek között a prefixumok csoportját nem tárgyalja, a grammatikai leírások ilyen morfématípust nem különítenek el. Elszórva, a nyelvi elemek leírásában azonban mégis találkozunk a *prefixum* megjelöléssel: a felsőfok *leg-* elemének és az igekötőknek a minősítésében.

A grammatikában, vagyis a szerkesztés tanában (tekhné grammatiké) jól körvonalazható a morfológia két fő területe: a morféma leírása, valamint a morféma összeszerkesztési szabályainak és e morfológiai műveletek eredményeként létrejött szerkezeteknek a vizsgálata. Eszerint a grammatikában (a szerkesztés tanában) a nyelvi egységeknek a következő szintjei figyelhetők meg:



Ha a lexéma–szóalak–szövegszó között lévő kapcsolatokat is be akarjuk építeni, a következő összefüggésrendszert kapjuk:



A szerkesztés alapegységei a morféma: szótövek, segédzők, affixumok. Összeszerkesztésükkel vagy újabb alapegységek – szerkesztett (tő)morfémák – vagy grammatikai (szó)alakok jönnek létre. Ez utóbbiak az alapelemei a glosszémáknak (szintagmatagoknak), azok pedig szintagmákká kapcsolódhatnak.

A *lexematizálás* feladata új lexémák elvonására alkalmas szótövek létrehozása. Főbb fajtái: a szóösszetétel és a szóképzés.

A *paradigmatizálás* paradigmata, azaz szóalakok megalkotására alkalmas. A kiinduló morféma itt is a szótő, ragozással szintetikus szóalakot kapunk (szótő+[jel+]rag), pl.: *gondol+t+am*; segédszóval pedig analitikusat

(szótő+[jel+]*rag* + *segédszó*): *gondol+t+am* + *volna*. A szótő kötelező, mellette a szerkesztésben zéró alakú morféma is részt vehet.

A **glosszematizálás** szintagmatag létrehozását jelenti. Ez teszi alkalmassá a szóalakot arra, hogy „felsőbb osztályba lépjen”, vagyis hogy szintagmatag szerepét betölthesse. Tágabb értelemben egy minősítő eljárásról van szó, amelynek során eldől, hogy a szóalak alkalmas-e glosszémaszerep betöltésére. Szűkebb értelemben a glosszematizálás ennek az eljárásnak a szerkesztési fázisa. Maga a szerkesztés szóalak és *segédszó* összekapcsolásával történik, eredményül a magyar nyelvben analitikus forma jön létre.

Minthogy egyes morfémák általában az egyik, mások a másik művelet-fajtának engedelmeskednek, a szóelemeket e műveletekkel is jellemezhetjük (csoportosíthatjuk).

Nyelvünk a fentiekkel ellentétes irányú változásokat is ismer. A *deszintagmatizálódás* során (szó)szerkezet olvad össze lexémává (pl. *cserben hagy* → *cserbenhagy*; ugyanígy: *jóhiszemű*, *légyott*, *mitugrász* stb.). A *deglosszematizálódásban* szintagmatag válik lexémává, pl.: *nap közti* → *napközi*; továbbá *földalatti*, *nemulass*, *nemtörődömség* stb. A *deparadigmatizálódás* folyamán szóalak közelít a lexémaértékhez, ilyenkor az új elem alakilag még nem illeszkedik az adott lexémaosztályba: pl. *lépte(i)m*, *lépte(i)d*, *lépte(i)* fn → *lépte* (‘lépése’) fn, illetve (az) *éljen(ek)*, (a) *félsz*, (a) *hiszem* ige → főnév. Ezeknek a változásoknak az a közös jellemzőjük, hogy mindegyik esetben új lexéma keletkezik, így összefoglaló néven *lexikalizálódásnak* nevezhetjük őket. (Hasonló, de más természetű jelenség a *delexikalizálódás*: szótő affixumértékű nyelvi elemmé válása. Így keletkezett pl. a *tő* főnévből a *-tól* ~ *-tól* határozórag; a *hat* igéből a *-hat* ~ *-het* képző stb.)

A fentiekből kitűnik, hogy a **prefixumot** mint kötött morfémat – ezt az elnevezése is jelzi – a szóalak vagy a szótő szerkezetében kell keresni. Tudniillik a prefixum a szóelemeknek az a csoportja, amely nem közvetlenül a glosszéma, hanem a lexéma vagy a szóalak létrehozásában vesz részt. Eszerint prefixum és *segédszó* funkcionálisan elhatárolódik egymástól. (Az utóbbit hasonló szórendi helyzetben *prepozíciónak* nevezzük.) A prefixum tehát a szóalak szerkezetében az a morféma, amely *a szótő előtt található, és a) új lexémát hoz létre, de szótóként nem kap nyelvi szerepet, vagy b) új szóalakot hoz létre, de c) glosszémaszerepre sem önmagában, sem segédszóval nem alkalmas.*

Ezek a tulajdonságok a magyar nyelvben a következő morfémacsoportokat jellemzik:

a) A magyar nyelvben az egyetlen paradigmatis prefixum a **felsőfokjel**. Funkciójában a jellel mutat rokonságot: megköti az alapszó szófaját, felsőfokjeles szóalakhoz képző nem járulhat; szintaktikai viszonyító szerepc

nincs. A *leg-* elsősorban grammatikai jelentés hordozója: a szótőben megjelölt tulajdonság legnagyobb mértékét, fokát fejezi ki.

b) A magyar nyelvbe – idegen szavakkal, azok részeként – bekerültek **idegen prefixumok** is (pl. *a-*, *de-*, *in-*, *re-* stb.). Szinte mind latin vagy görög eredetű, a nyelvérzék annyira idegen elemeknek tartja őket, hogy magyar lexémákhoz nem is kapcsolja, pl. *decentralizál*, *indiszkkrét*, *posztpozíció*, *unszimpatia* stb. Kivételesen néhány magyar szótő előtt is megjelennek, pl. *antitest*, *antianyag*, *exelnök*, *posztkárári*, *ultrabal* stb.

c) Az **igekötő** több tulajdonsága is a képzőkkel rokonítja. Az igekötőknek nincs állandó alaki önállóságuk, szövegszóként csak bizonyos típusú szerkezetekben (pl. tagadás) fordulhatnak elő. De ezt az alaki önállóságot még szófaji tulajdonságok is korlátozzák (az igekötős névszók nem válnak el: *elbeszélés*, *feladat*, *megfelelő*). Az igekötő a mondatban lényegében nem önállóan szerepel, mindig egy szóalakból lép ki (*megvár* → *meg sem vár* → *ne várd meg!*). Igekötő és ige(név) kapcsolata mindig morfológiai, mondatfunkciós képessége nincs. Szótővel összekapcsolódva nem a szintagma-, hanem a szóalkotás eszköze. A magyar igekötő korlátozott alaki önállósága a lexémaértéknek nem bizonyítéka, hanem történeti hagyatéka. Eszerint tehát az igekötő (derivációs) prefixum. (Részletesebben lásd a 2.2. pontban.)

d) Szigorúan leíró szempontból nézve a magyar nyelvben sok olyan alakulatot találunk, amely szerkezetét tekintve összetett szó, de előtagja az adott formában, jelentésben, szófaji értékben önállóan nem használatos. Az ide tartozó elemek nagy szóródást mutatnak a prefixum és a (fiktív) összetételi előtag között. Az utóbbitól haladnak az előbbi felé. Jellemzőjük, hogy szemantikailag önállóak, alaktanilag azonban kötött szóelemek, önálló lexémaként más értékben szerepelnek, mint „előtagként” (vö. B. Lőrinczi 1970, Berrár 1975 és 1983). Így szófaji értékük sem határozható meg pontosan. Összefoglalóan **prefixoidoknak** nevezhetjük ezeket az önállótlan morfémákat. Közöttük vannak **prefixumszerű előtagok**, amelyek ugyan előtagértékűek, de alaktani kötöttségük az affixumokhoz közelíti őket, pl.: *egyen- (-jogú, -ruha, -súly stb.)*, *mellék- (-épület, -folyó)*, *röp- (-labda, -pálya, -tér)* stb. Ha azonban a gyakoriság társul az előbb részletezett alaktani kötöttséggel, akkor az előtagszerű érték többnyire másodlagossá válik. Ebben az esetben már inkább **előtagszerű prefixumokról** (*al-*, *ellen-*, *magán-* stb.) beszélhetünk. A **valódi prefixumokban** e két tulajdonság mellett lényeges a harmadik, a jelentésbeli is. Ennek megítélésében már eléggé ingoványos területre jutunk. A jelentések módosulásának mértékét nehéz meghatározni. Mégis úgy tűnik, hogy az *elő-*, *fő-*, *köz-*, *mű-*, *ön-* elemek nagy produktivitása nem véletlen, a jelentésszerkezetben is kimutatható a módosulás. Ezekben a funkcióváltás olyan előrehaladott állapotú, hogy prefixumoknak minősíthetjük őket.

A *prefixum* terminusnak általánosan elfogadott magyar megfelelője nincs. A funkció és a pozíció lényegi jellemzőit egyaránt jól érzékelteti Tompa József *előtoldalék* megnevezése (1961: I, 528), s ezt veszi át Velcsovne is (1968: 189). Egyszerűsége, tömörsége miatt szerencsés Deme László *előtét* kifejezése (1976: 44).

2.2. Az igekötők vizsgálata

A magyar igekötő kutatásának gazdag szakirodalma van. Jakab István nemrég külön tanulmányban részletezte, hogy a különböző időkben – Sylvester Jánostól napjainkig – ki minek tartotta az igekötőt, és hogy ki mit tartott igekötőnek (Jakab 1982). Egy korábbi munkájában pedig a meghatározás és az állománykijelölés kérdéseit taglalta (Jakab 1976).

Az igekötőkről máig a legátfogóbb, legalaposabb munkát J. Soltész Katalin készítette (1959). Mértékadó monográfiája óta a legjobb szintézist D. Máta Mária adta (1989). A vizsgálat történeti szempontú, ám korszerű szemlélete folytán a leíró elemzés számára is fontos eredményeket ad.

Új szempontok és módszerek függvényében kap minősítést az igekötő a generatív nyelvelírásban mind a mondat szerkezet, mind az aspektus kutatásában. (Lásd pl. É. Kiss Katalin, Kiefer Ferenc és Wacha Balázs tanulmányait.)

Az igekötők nyelvünknek nem ősi örökségei. A különböző finn nyelvekben nem találjuk nyomát, a két obi-ugor nyelvben viszont a magyarhoz hasonló gazdag igekötőrendszer fejlődött ki. Az igekötőt a kutatók olyan átmeneti jelenségnek tartják, amely lativusi határozószóból különböző fokozatokon át a képzői érték felé tart. A változás megindulása felől az igekötő többé-kevésbé *szóértékű*; a folyamat egészét tekintve *háromarcú* jelenség: eleinte szó, később önállóságát veszített kötött helyzetű előtag, végül funkcióváltással képzőértékű morféma. A végpont felől nézve: új lexémák létrehozására alkalmas *képző*. A céltól függően mindegyik szempontnak és módszernek lehet létjogosultsága. Ebben az értekezésben a harmadik, a szinkroniát erőteljesebben érvényesítő megközelítés érvényesül, vállalva e kevésbé elfogadott álláspont kritikus megítélését.

A formális szempontú elemzésben szükséges az igekötőt tartalmazó alakulatok morfológiai és szófaji tulajdonságainak áttekintése. Az elemzésben sarkalatos kérdés az elváló igekötők viselkedési szabályainak explicit leírása is. A dolgozat nem tér ki olyan lényeges, ám a formális elemzés számára irreleváns tulajdonságok taglalására, mint a történeti fejlődés, az aspektus vagy a jelentés kérdése. Ezeket a területeket sem zárja ki a vizsgálatból, de jellegzetesen nyelvtörténeti, dialektológiai stb. kérdésekkel, azok kimerítő leírásával nem foglalkozik.

Az igekötő elsődleges funkciója a *lexematizálás*, vagyis új lexéma elvonasára alkalmas tömorféma létrehozása. Ez természetszerűleg mindig a szimplex ige valamilyen *jelentésmódosulásával* jár együtt (speciális vagy irányjelentés, aspektualitás, cselekvésminőség stb. kifejezése). A jelentésszerkezetben történő változás többnyire kihat az ige *valenciájára* is, ami a vonzatkeret megváltozásában nyilvánul meg. Nem szófajváltó és nem szófajkötő elem, a (szuffixális) továbbképzés lehetőségeit meghagyja.

Az igealakok szerkezetében a szótövet megelőző morfémák a következő csoportokba sorolhatók:

a) A *valódi igekötők* gyakoriságuk és szemantikai tulajdonságaik alapján tisztán prefixumértékűek, pl. *át, be, el, elő, meg, oda, szét* stb.

b) Az *előtagszerű igekötők* még sokat megőriztek határozószói értékükből: *abba, elé, félre, közbe, neki, ott, rajta, széjjel* stb. (Vö. Jakab 1976: 98–99.)

c) A magyar igékre nem jellemző a szóösszetétel. A kevés számú ilyen alakulat egy részében az *igekötőszerű előtagok* (igei prefixoidok, pl.: *cserbenhagy, egyetért, jólesik, síkraszáll* stb.) az igekötőkhöz hasonlóan viselkednek: elválnak igéjüktől.

d) A *valódi előtagok* szilárd összetételt alkotnak az igével: *mennydörög, pártfogol, számfejt* stb., továbbá a nyelvhelyességi szempontból vitás *műfordít, gépir, szakfelügyel* (sőt újabban: *kisvállalkozik*).

Sokat vitatott téma az *igekötők szórendje*. Nyelvtanaink általában elfogadják az igekötő szórendjének formai alapú hármas osztályozását (egyenes, megszakított, fordított). A formális elemzés következetes alkalmazásával azonban e rendszerezés több ponton vitatható.

E felosztás szerint *egyenes szórendűnek* nevezzük az igekötőt akkor, ha az ige(név) előtt áll, azzal összeforrv. Csakhogy az ilyen igekötő egyetlen szövegszót – és következésképpen: szóalakot – alkot igéjével, s így esetükben valójában nem szórendi, hanem a morfémák *sorrendi* tulajdonságairól van szó. Ha következetesen akarunk lenni, azt kell mondanunk, hogy szórendjük voltaképpen csak az elváló igekötőknek van. Ez viszont újabb nehézséget támaszt. A nem elváló igekötők kihagyása a további elemzésből módszertanilag ugyan helyes, rendszertanilag azonban nem fogadható el, hisz mellőzésükkel kizárnánk a vizsgálatból az igekötős szóelőfordulások csaknem 4/5 részét. Az ellentmondást úgy oldhatjuk föl, hogy a továbbiakban nem szórendi, hanem az igekötős ige(név) elemeinek egymáshoz viszonyított *sorrendi* jellemzőiről beszélünk. (Az elválók esetében természetesen a *szórend* megnevezés is helytálló.)

Az ún. *megszakított szórendi* típusnak nem is elsődlegesen az elnevezése, inkább a besorolása kifogásolható. Tudniillik az ilyen sorrendű igekötős ige (*meg fogja mondani*) is egyenes „szórendű”, hisz a kettő nem zárja ki egy-

mást. A megszakított sorrendű szerkezet nem az egyenes mellett létezik, hanem benne. Vagyis ez nem fő-, hanem altípus.

Ugyancsak megfontolandó tény az is, hogy beékelődés nemcsak egyenes sorrend esetén lehetséges: jóval ritkábban ugyan, de a **fordított szórendű** szerkezetekbe is ékelődhetnek be más nyelvi elemek. Ha ezt is önálló változatnak tekintjük, akkor kiderül, hogy *a megszakított szórendként tárgyalt jelenség nem egyetlen egységes, az előzőekkel párhuzamos, hanem azokat keresztező típus.*

	kapcsolt	megszakított
egyeses	<i>megkérdezi</i>	<i>meg is kérdezi</i>
fordított	<i>kérdezd meg!</i>	<i>kérdezd csak meg!</i>

Ezek figyelembevételével az igekötős szóalakok **sorrendi** (és nem szórendi!) **típusait** a következőképpen vázolhatjuk fel:

1. egyeses sorrend: a/ (össze)kapcsolt: *fölad*;
b/ megszakított: *föl is adja*;
2. fordított sorrend: a/ kapcsolt: *add föl*;
b/ megszakított: *add már föl*.

2.3. Automatikus elemzés

Ebben a részben történik – az előző fejezet elemzései alapján – az igekötős alakulatok automatikus szintézisének és analízisének leírása. **Szintézisen** itt nem szintetikus eljárásmódot értünk, hanem a szövegben analitikusan (széttagoltan) álló igekötős és összetett igék, igenevek reprodukálását, a kiinduló alakok helyreállítását. És fordítva, **analízisen** pedig az előképzős és összetett lexémák előkészítését különféle nyelvi műveletek számára.

A magyarban az igekötős szóalakokból tehát 3 **analitikus** szintaktikai-morfológiai szóalak vagy szerkezet hozható létre. Közöttük leggyakoribb a kapcsolt fordított (*nem tudja meg*), legbonyolultabb a megszakított egyeses (*Laci meg akarta még azt is tudni*), kevésbé jelentős a megszakított fordított szórendű szerkezet (*nem jöttek volna rá*). Ezek általában **igék** (*el is adta*), de lehetnek **igenevek** (*ki kell azt is vinned, el voltak már fáradva, nem adható ki*), ritkán **főnevek** is (*hozzá nem értés*). A továbbiakban a fenti három típust – amelyek tehát az igekötőn és a hozzá tartozó igén vagy igenéven kívül beékelődő elemet is tartalmazhatnak – összefoglaló néven igekötős szerkezeteknek fogjuk nevezni.

Összegezve, a beékelődő elemek és sorrendjük a következő:

1. hely: **igekötő** (PR)

2. hely: *is, ne, nem, se, sem, úgysem, mégsem* (=IS)

3. hely: a) *kell, lehet, akar, szeret, tud, bír, mer, óhajt, talál; fog, szokott*
(KELL típus)

b) *szabad, képes, muszáj* (SZABAD)

c) *van, volt, lesz, nincs, sincs* (VAN)

4. hely: *volna*

5. hely: *még, már, majd, most, megint* (MÉG)

6. hely: *az, ez, azt, ezt, arra, erre, azon, ezen* (AZT)

7. hely: *is (sem)*

8. hely: ? (nem tipizálható)

9. hely: a) **ige** (V)

b) **igenév** (NV)

A beékelődés jellege és bonyolultsága szófajspecifikus. Finitív igébe kizárólag IS csoportbeli elem ékelődhet be (pl.: *le sem ült*). Az igenevek közül az infinitívuszos szerkezetben akár mind a 6 (illetve 7) hely kitölthető (*ki sem tudta volna még azt sem találni*), adverbiumban az ékszők halmozása jóval ritkább (*be is volt zárva; ha be lett volna már az is zárva*). A particípiumok közül az állítmányi szerepű hatóképzősekbe IS ékelődhet (*el sem lehet, meg is indokolható*), egyéb esetekben csak a *nem* tagadószó fordul elő, az *is* igen ritkán (*a soha el nem szenderedő természet, vissza nem utasított igények*). Névszót csak kivételesen bont meg éksző (*meg nem értés*).

A beékelődés lehet **egyszerű** (egyetlen szóalak, pl.: bele tud nyugodni, neki ne menj), **szerkesztett** (analitikus beékelődéssel: elő lehetett volna adni) és **kombinált** (több típusból összekapcsolva: be sem tudtak lépni, rá kellett volna még azt is írni). A gyakorlatban azonban igen ritka az ilyen bonyolult szerkesztés. A vizsgált mintegy 150 000 szövegszós korpusz megszakított szerkezeteinek csupán az 5,47%-ában fordul elő egynél több éksző. Közülük az egyik mindig KELL típusú ige, az esetek felében a *kell*, a másik felében (gyakoriságuk rendjében) a következők: *akar, lehet, tud, szeret(ne), szokott*. Előttük csak IS állhat (vissza is kellett jönni; el sem tudja képzelni stb.). A KELL utániak között leggyakoribb a *volna* (haza szeretett volna menni), azt követi a *már* (el kellene már döntenem). A KELL után nem tipizálható bővítmény is állhat (pl.: be kell pótlólag vezetnünk az x operátort; mert be akartunk a színházba érni).

Igekötőt tartalmazó három szerkezettípusunk tehát: a) igekötő (PR) + beékelődés + ige (V) vagy igenév (NV); b) ige + igekötő; c) ige + beékelődés +

igekötő. Az a) jelű szerkezet beékelődése 1–6 elemből áll, melyek meghatározott rendben követik egymást:

PR + IS + KELL + volna + MÉG + AZT + is + V/NV,

A c) típus szerkezete egyszerűbb: V + volna/csak/MÉG/AZT + is + PR.

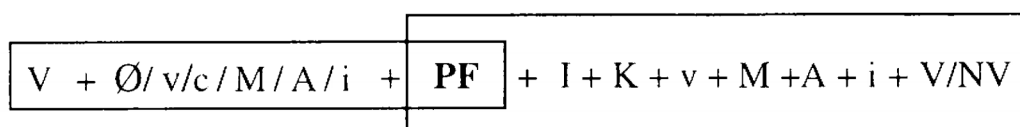
A jelzéseket leegyszerűsítve a három szerkezetet így írhatjuk fel:

a/ PR + [I + K + v + M + A + i] + V/NV

b/ V + Ø + PR

c/ V + [v/c/M/A + is] + PR.

Ha a három típust egymásra vetítjük, olyan kváziszerkezetet kapunk, amely mindenféle igekötős alakulat modellálására és elemzésére alkalmas:



A kapott álszerkezetet úgy értelmezhetjük, mint különböző tengelyek mentén elhelyezkedő részhalmazok metszeteinek rendezett halmaza. Ez utóbbi adja meg a *szintagmatikus tengelyt*, amelynek tagjai az igekötős szerkezetek elemcsoportjainak változói. A *paradigmatikus tengelyek* mentén pedig található az igekötős szerkezetek – a halmazelmélet kifejezésével élve – igazsághalmazainak elemei. E részhalmazoknak természetesen üres elemei – és így üres metszetei – is vannak.

Visszatérve eredeti feladatunkhoz, az elváló igekötő igéjével vagy igenevével való automatikus összekapcsolásához az igekötőt mint a szintagmatikus tengely első azonosított elemét úgy foghatjuk föl, hogy az egy (kvázi)-szerkezet közepén áll. A jobbra keresés beékelődő egyenes szórendet tételez föl. A lépések a következők:

1. E fázis első lépése, az igekötő utáni 1–2. pozíció vizsgálata perdöntő, cz minősíti a szerkezetet. A PR utáni 1. helyen csak IS, a 2.-on pedig KELL típusú ékszó fordulhat elő. Három eset lehetséges:

1.1. Igekötős ígéről van szó, a beékelődés kizárólag az IS tömb valamelyik tagja lehet. Ebben az esetben a KELL típus le van tiltva (pontosabban a tömböt kötelezően a 0. elem képviseli), de a többi – minthogy valamennyi KELL-függő – előfordulása is ki van zárva:

PR + IS+0 + V → PR-V + IS (V = KELL),

azaz: ha az IS+1. elem ≠ KELL, akkor igekötős ígét találtunk, a konkatenációt el kell végeznünk. Pl.: *rá sem nézett* → *ránézett* + *sem*. A kimeneten megkaptuk a keresett lexémát, az adatot tároljuk vagy továbbítjuk, mindenestre a szintézis művelete befejeződött.

1.2. A 2. helyen KELL típusú ékszó áll. Ennek a legnagyobb a hatása a szerkezetben: törli az IS minősítő funkcióját, determinálja az összes további

műveletet, jelzi a szerkezet igenévi értékét – az őt megelőző és az utána következő elemek minőségétől függetlenül.

Három eset lehetséges: a) a KELL után nincs több ékszó, az összekapcsolás elvégezhető, a kimeneten infinitivust kapunk; b) a KELL után van ékszó, ezeket sorban azonosítani kell, az ékszók után főnévi igenevet kapunk; c) a tömböt annak egyik altömbje reprezentálja, melynek elemei a létige szuppletív alakjainak egyike: adverbiumot kapunk. (A létigét ritkán MÉG követi.) A műveletek:

a/ $PR + KELL + INF \rightarrow PR-INF + KELL$

b/ $PR + KELL + volna + AZT + is + INF \rightarrow PR-INF + \text{egyéb}$

c/ $PR + VAN (+ MÉG) + ADV \rightarrow PR-ADV + VAN$

A leírt műveleteket biztosabbá lehet tenni, ha összekötjük a megtalált INF vagy ADV végződéseinek az ellenőrzésével. Az ADV esetében ez egyszerű, az INF toldalékos alakjainak rendszere is könnyen áttekinthető. (Vö. Antal 1964: 11 kk. A verbum finitum ilyenfajta leírására eddig egyetlen kísérlet született: Jánoska 1967: 464–8.)

1.3. Amennyiben mindkét preferált tömbünk 0 fokon van jelen, vagyis

$PR + 0 + 0 + \dots$,

akkor keresésünk eredménytelennek bizonyult, vissza kell térnünk az igekötőhöz, hogy a művelet 2. fázisát elvégezhessük.

2. Amennyiben a halmazok metszeteinek azonosítása nem járt eredménnyel, a keresést a másik irányban kell folytatnunk. A keresési fázisok ilyen sorrendjével annyit most biztosan tudunk, hogy az önállóan szereplő igekötő nem megszakított, hanem fordított szórendű szerkezet tagja. Ezért, valamint a beékelődés sokkal áttekinthetőbb volta miatt az elemzés jóval egyszerűbb. Csupán az igekötő előtti szót kell megvizsgálni.

2.1. Az igekötő előtti pozícióban csak az IS tömb *is* eleme (*hozd is be, mondd azt is el*), ezt megelőzően pedig a *csak* (*ülj csak le*), ritkábban az AZT vagy a MÉG tömb valamelyik eleme (*tedd már le*) fordulhat elő:

$V + volna/csak/MÉG/AZT+(is) + PR \rightarrow PR-V + \text{egyéb.}$

2.2. A leggyakoribb eset azonban az, hogy az ige és az igekötő között nem áll ékszó, az összekapcsolás elvégezhető:

$V + 0 + 0 + PR \rightarrow PR-V.$

A számítógépes nyelvészet és a gépi előtt még óriási feladatok vannak. Az itt leírtak ennek csupán egy kis részletét tisztázzák. De arra jó, hogy rámutasson: megoldhatatlannak tűnő feladataink megfejtéséhez sokszor valami gyökeresen más ráközelítési mód, új szemlélet visz bennünket közel.

3. A vizsgálati eredmények hasznosítási lehetőségei

1. A dolgozat első részének megállapításait elsősorban az elméleti nyelvtudomány – morfológia, szófajtan, szórendi kutatások – tudja felhasználni.

a) Nyelvünk taxonomikus morfológiai leírását ki kellene egészíteni a prefixumok típusával. Fogalmának meghatározásához – más nyelvek leírása és hazai próbálkozások felhasználásával –, valamint rendszerezéséhez javaslatot adtam.

b) Az igekötők funkcióinak újragondolása módosításra készíti a szófaji felosztást is. Ha ugyanis az igekötő nem szóértékű, akkor nincs helye a szófaji rendszerben.

c) Felül kell bírálni az igekötős szerkezetek szórendi tipizálását is. Mint kifejtettem, nem 3, hanem 2×2 típussal kell számolnunk.

2. A megállapítások alkalmazhatók (vagy alkalmazandók) az oktatásban.

a) Az igekötőkkel kapcsolatos megállapítások egy része (pl. a szórendi típusok) hasznosíthatók az általános és középiskolai anyanyelvoktatásban.

b) Úgy gondolom, az egyetemi-főiskolai magyar nyelvészeti oktatásba is érdemes lenne bevezetni a fölvázolt morfológiai rendszerezést, valamint az igekötők tárgyalásának javasolt módszerét.

c) A magyarnak mint idegen nyelvnek a tanításában a nehézségek között szerepel az igekötős igék, igeelvek használatának elsajátíttatása. Saját tapasztalatom szerint különösen az igekötő szórendjének tanításában alkalmazhatók eredményesen az itt bemutatott módszerek.

3. Az igekötő formai clemzésének (leválasztás és összekapcsolás) fontos felhasználási lehetőségét a különböző gépi feldolgozások jelentik. Itt a következő alkalmazási területek látszanak:

a) A kvantitatív módszerű kutatásokban tarthatatlan az a hibás gyakorlat, miszerint szófaj-statisztikai vizsgálatokban csak az elváló igekötőket regisztrálják. Eszerint ami elválik, az igekötő, ami nem, az meg nincs is, illetve az adott igével, igenévvel, névszóval együtt egy szövegszó. A dolgozat mind az igekötő, mind az igekötős alakulatok vizsgálatára módszertani javaslatot ad.

b) Arra is találunk példát, hogyan készíthetők konkordencialistából viszonylag egyszerűen – és ami fontos: gépi úton – más (célú vizsgálatoknak megfelelő) listák is.

c) Az említett nehézségek jelentkeznek a gépi szótárkészítésben is. Ma már igen jelentős terjedelmű szövegek vihetők gépre, s nyerhetők a célnak megfelelően a legkülönbözőbb típusú szótárak (betűrendes, gyakorisági, a

tergo stb.). A lemmatizálás során itt is elkerülhetetlen az igekötős alakok reprodukálása (egyeses szórendűvé alakítása).⁴

d) Napjainkban szerte a világon rendkívül elterjedtek a szövegszerkesztő programok. Ma már egy intelligens szövegszerkesztő nem nélkülözheti az ún. helyesírás-ellenőrző (spelling checker), valamint automatikus szóelválasztó programot. Ezek sajnos ma még csak „angolul tudnak”, de vannak már piacképes magyar szoftverek is. Sajnos az igekötővel még egyik sem tud bánni. (Emiatt találunk ilyen elválasztásokat: *me-gáll, előz-etes, elé-nekel, fel-elet, a fel-ének* stb.)

e) Végül igényli e kis részprobléma megnyugtató rendezését a gépi fordítás is. Ez azonban újabb, szerteágazó kérdéscsoportot vet föl. Pl. az igekötő viselkedésének leírásához másképp kell nekiindulni, ha magyarra, s másképp, ha magyarról fordítunk. De ez már egy másik dolgozat témája lehetne.

Az értekezés témakörével kapcsolatos publikációk

Igekötős szerkezeteink szórendi jellemzői. *Tanulmányok a magyar nyelvről*. TK. XVIII, 107–120. Eger, 1987.

Egy ritkán tárgyalt morfémafajtánkról. *Magyar Nyelv* LXXXIV, 473–476. (1988)

Igekötős alakulataink szófaj-gyakorisági jellemzői. *Tanulmányok a magyar nyelvről*. TK. XIX./II, 3–11. (1990)

Igekötős szerkezetek automatikus elemzése. In: *I. Magyar Alkalmazott Nyelvészeti K.* (Szerk.: Székely Gábor) Nyíregyháza, 1991. 699–703.

Igekötős szóalakok morfológiai elemzése számítógéppel. *Tanulmányok a magyar nyelvről*. TK. XX, 7–12. (1991)

A magyar igekötők szórendi típusai. *Magyar Nyelv*, LXXXIX, 312–318. (1993)

A magyar szórend egy sajátos jelensége. *Tanulmányok a magyar nyelvről*. TK. XXI, 33–39. (1993)

Az igekötők használatának tanítása idegen ajkúak számára. In: *III. Magyar Alkalmazott Nyelvészeti K.* (Szerk.: Klaudy Kinga.) Miskolc, 1993. II, 279–284.

A gyakorisági listák egy hibatípusáról. *Magyar Nyelvőr* 118: 43–46. (1994)

⁴ „Egybeírásnál a szótárbeli elemekkel való összevetés nem jelent problémát. Különírásnál azonban nem tudunk olyan eljárást, mely egy mondatbeli potenciális igekötő (például *oda, be* stb.) és egy ige együvé tartozását megerősítené vagy elvetné. Ezért az egyetlen elfogadható eljárásnak a felhasználó megkérdezését tartjuk” (Prószéki 1989: 534).

- Hányféle szórendje van a magyar igekötőnek? *Magyartanítás*, 1994. 5, 25–26.
- A prefixumok helye a morféimák rendszerében. *Tanulmányok a magyar nyelvről*. TK. XXII, 5–14. (1994)
- Az igekötős alakulatok formai vizsgálata. *Tanulmányok a magyar nyelvről*. TK. XXIII, 142–156. (1997)
- Gondolatok a magyar igekötők természetéről. *Tanulmányok a magyar nyelvről*. TK. XXIV, 262–276. (1998)
- Bevezetés a magyar nyelv morfológiájába*. (Úvod do morfológie maďarského jazyka) Egyetemi jegyzet. Banská Bystrica, 1999.
- A szavak. In: *Ismertek a nyelvről*. Hungarovox Kiadó, Budapest, 2001. 123–143.

Hivatkozások

- Antal László 1961. A magyar esctrendszer. *Nyelvtudományi Értekezések* 29. sz.
- Antal László 1964. *A formális nyelvi elemzés*. Gondolat Kiadó, Budapest.
- Antal László – Csongor Barnabás – Fodor István 1970. *A világ nyelvei*. Gondolat Kiadó, Budapest.
- Berrár Jolán 1974. Új szempontok és módszerek a szóképzés vizsgálatában. In: *Tanulmányok a mai magyar nyelv szófajtana és alaktana köréből*. Tankönyvkiadó, Budapest. 99–122.
- Berrár Jolán 1975. Morfológiai szerkezetek – szintaktikai szerkezetek. *Magyar Nyelv* LXXI, 35–40.
- Berrár Jolán 1983. Próbák és problémák a Mai magyar nyelv c. tankönyv új kiadásához. *Nyelvtudományi Dolgozatok* 33. Kézirat.
- Deme László 1976. *A beszéd és a nyelv*. Tankönyvkiadó, Budapest.
- Fleischer, Wolfgang 1971. *Wortbildung der Deutschen Gegenwartssprache*. VEB Bibliographisches Institut, Leipzig.
- Gaál Edit 1991. Egy magyar nyelvtan terve. *A mai magyar nyelv*. (Cikk- és tanulmánygyűjtemény) Összeállította: Keszler Borbála. Tankönyvkiadó.
- Jakab István 1976. A magyar igekötők állományi vizsgálata. *Nyelvtudományi Értekezések* 91. sz.
- Jakab István 1982. *A magyar igekötő szófajtani útja*. *Nyelvtudományi Értekezések* 112. sz.
- Jánoska Sándor 1967. A magyar ige automatikus toldalékolásának egy modellje. *Nyelvtudományi Értekezések* 58, 464–468.
- Károly Sándor 1980. Szavak, szerkezetek morfológiai tagoltságának változásai és a jelentés. In: *Tanulmányok a mai magyar nyelv szóképzéstudana és jelentéstudana köréből*. Tankönyvkiadó, Budapest. 121–156.
- B. Lőrinczy Éva 1970. Szintaktikailag pontosan nem elemezhető összetett szavainkról. *Magyar Nyelv* LXVI, 63–75.
- D. Mátaí Mária 1989. Igekötőrendszerünk történetéből. *Magyar Nyelv* LXXXV, 9–28, 151–171., ill. MNYTK. 187. sz.

- Papp Ferenc 1975. *A magyar főnév paradigmatisztikus rendszere*. Akadémiai Kiadó, Budapest.
- J. Soltész Katalin 1959. *Az ősi magyar igekötők*. Akadémiai Kiadó, Budapest.
- T. Somogyi Magda 1992. Toldalékaink. In: *Újabb fejezetek a leíró magyar nyelvtan köréből*. (Szerk.: Keszler B.) Tankönyvkiadó, Budapest. 225–232.